



**PEMBUATAN POHON BERBOBOT UNTUK PENCARIAN SEMANTIK
MENGUNAKAN ALGORITMA *WEIGHTED TREE SIMILARITY* PADA
PENILAIAN DOKUMEN RENCANA PERKULIAHAN
(SAP, GBPP/SILABUS DAN KONTRAK KULIAH)**

Febi Eka Febriansyah¹⁾ dan Astria Hijriani¹⁾

¹⁾Dosen Jurusan Ilmu Komputer FMIPA Universitas Lampung
Jl. Prof. Dr. Soemantri Brodjonegoro No. 1 Bandar Lampung 35145
Surel: febi.febriansyah@fmipa.unila.ac.id

ABSTRACT

In order to speed up the examination process of the lecture plan documents (GBPP, SAP and Contracts Lecture), needed an application that is able to match the format of writing the document. Semantic search method with metadata, one of which is weighted tree similarity, it is very suitable to be applied in this application. This study focused on the establishment of weighing tree of weighted tree similarity algorithm for searching and matching documents lecture plan that matches the format of the assessment standards PPMU. Matching only aimed at the main key words, such as GBPP, SAP, Course, Department, Course Descriptions, Competence, etc. A comparison is made with tree branches without a label, only the value of the weight on the branch. The amount of weight on each branch is divided equally by the number of branches, its value between 0 to 1.

Keywords: Contracts Lecture, GBPP, SAP, Semantik Search, *Weighted Tree Similarity*.

ABSTRAK

Guna mempercepat proses pemeriksaan dokumen rencana perkuliahan (GBPP, SAP dan Kontrak Kuliah), dibutuhkan sebuah aplikasi yang mampu mencocokkan format penulisan dokumen tersebut. Metode pencarian semantik dengan metadata, salah satunya adalah *weighted tree similarity*, sangat cocok diterapkan pada aplikasi ini. Penelitian ini dipusatkan pada pembentukan pohon (*tree*) berbobot dari algoritma *weighted tree similarity* untuk pencarian dan pencocokan dokumen rencana perkuliahan yang sesuai dengan format standar penilaian dari PPMU. Pencocokan hanya ditujukan pada kata-kata kunci utama, seperti GBPP, SAP, Matakuliah, Jurusan, Deskripsi Matakuliah, Kompetensi Dasar, dsb. Pohon perbandingan dibuat dengan cabang tanpa label, hanya ada nilai bobot pada cabang. Besarnya nilai bobot pada tiap cabang dibagi sama rata berdasarkan jumlah cabang, nilainya antara 0 (nol) sampai dengan 1 (satu).

Kata kunci: GBPP, Kontrak Kuliah, Pencarian Semantik, SAP, *Weighted Tree Similarity*.



PENDAHULUAN

Latar Belakang

Dalam perencanaan proses perkuliahan, dokumen rencana perkuliahan merupakan kelengkapan yang harus diserahkan oleh dosen kepada jurusan pada pekan awal dimulainya semester ganjil/genap (Prosedur ISO9001:2008 PM-9, 2014). Selama ini dokumen diserahkan dalam bentuk *hardcopy* dengan muatan sesuai mata kuliah yang diampu setiap dosen. Dokumen akan diverifikasi Tim Penjamin Mutu Program Studi (TPMPS). Arsip dokumen perkuliahan akan diperiksa oleh tim audit dari Pusat Penjamin Mutu Universitas (PPMU) pada setiap akhir tahun ajaran.

Proses verifikasi yang dilakukan oleh TPMPS saat ini, hanya berupa proses pemeriksaan kelengkapan dokumen, ada atau tidaknya dokumen Satuan Acara Pembelajaran (SAP), Garis-Garis Besar Program Pembelajaran (GBPP) dan kontrak kuliah. Sedangkan hal yang menjadi perhatian dari audit PPMU jauh lebih detail, yaitu isi dari dokumen perkuliahan (Borang Monev Proses Perkuliahan Bermutu SOPMO2: 2013). Selain penyimpanan *hardcopy* dokumen perkuliahan yang kurang efisien karena membutuhkan ruang fisik yang tidak kecil, masalah juga muncul jika TPMPS harus memeriksa secara manual seluruh dokumen rencana perkuliahan sesuai standar operasional dari PPMU. Ada kebutuhan untuk proses pencarian informasi yang menelusuri dokumen yang dikumpulkan secara digital dan pencocokan akan terpenuhinya standar dari PPMU atas masing-masing komponen, baik SAP, GBPP dan kontrak perkuliahan.

Proses *matching* didahului oleh proses pencarian dokumen berbasis teks. Struktur penilaian dokumen rencana perkuliahan yang memperhatikan komponen dari SAP, GBPP, dan kontrak kuliah seperti identitas, standar kompetensi, alokasi waktu,

buku sumber, akan lebih tepat menggunakan pencarian semantik. Setiap komponen akan dianggap sebagai *query* kepada sistem dengan struktur tertentu. Struktur yang lengkap akan mendapatkan kemiripan yang lebih baik yang sesuai dengan standar penilaian dokumen perkuliahan.

Salah satu metode yang dapat digunakan untuk meningkatkan relevansi dokumen rencana perkuliahan yang sesuai dengan standar dari penilaian PPMU adalah dengan menyusun metadata yang lebih lengkap dan operasi pencocokan semantik menggunakan metode *weighted tree similarity* (Rahutomo, 2008).

Penelitian ini dipusatkan pada pembentukan pohon (*tree*) berbobot dari algoritma *weighted tree similarity* untuk pencarian dan pencocokan dokumen rencana perkuliahan yang sesuai dengan format standar penilaian dari PPMU. Pencocokan hanya ditujukan pada kata-kata kunci utama, seperti GBPP, SAP, Matakuliah, Jurusan, Deskripsi Matakuliah, Kompetensi Dasar, dsb.

SAP, GBPP dan Kontrak Perkuliahan

SAP (Satuan Acara Perkuliahan) adalah rencana kerja kegiatan perkuliahan setiap pertemuan yang mencakup tujuan umum, tujuan khusus, pokok bahasan, sub pokok bahasan, kegiatan belajar mengajar (tahap, kegiatan pengajar, kegiatan mahasiswa, media dan alat pengajaran) per pertemuan.

GBPP adalah pedoman pokok-pokok pembelajaran yang disusun secara sistematis yang mencakup terdiri dari deskripsi singkat mata kuliah, diagram analisis instruksional mata kuliah, kompetensi dasar, pokok bahasan, metode, media/alat bantu, waktu, daftar pustaka per minggu.

Kontrak Perkuliahan terdiri dari prasyarat, deskripsi singkat, tujuan, strategi perkuliahan, bahan bacaan (buku teks wajib dan buku pendukung), evaluasi (kuis,



tugas terstruktur, ujian praktikum untuk mata kuliah yang ada praktikumnya, ujian tengah semester/UTS, dan ujian akhir semester/UAS), dan sistem penilaian, Jadwal perkuliahan (pertemuan, tanggal, dan waktu kuliah) per pertemuan.

Pencarian Semantik

Semantik adalah studi tentang makna dalam komunikasi (Oxford,1994). Kata tersebut berasal dari bahasa Yunani. Di dalam linguistik berarti studi interpretasi tanda di dalam keadaan dan konteks tertentu (Wikipedia, 2015). Singkatnya semantik berarti makna sesuatu konten berdasarkan konteks tertentu. Suatu konten dapat memiliki banyak makna tergantung konteks ia berada. Sebagai contoh pakaian hitam sebagai konten. Di dalam konteks acara kematian, pakaian hitam bermakna duka cita, sedangkan di dalam konteks acara formal, pakaian hitam bermakna formalitas dan kewibawaan. Di dalam contoh ini pakaian hitam adalah konten sedangkan jenis acara sebagai konteks.

Apabila semantik berarti makna sesuatu konten berdasar konteks tertentu dengan demikian maka pencarian semantik adalah pencarian suatu konten berdasarkan konteks yang tepat. Mungkin terdapat konten yang sama dalam suatu sumber informasi tetapi memiliki konteks yang berbeda-beda. Tujuan dari pencarian ini agar diperoleh artikel yang relevan sesuai dengan konteks yang dimaksud pengguna. Dengan demikian untuk mencari dokumen yang relevan dalam pencarian semantik, pengguna harus memberikan *query* yang lebih lengkap pada sistem, baik konten maupun konteksnya.

Dalam pencarian artikel secara semantik, konten ditetapkan sebagai teks yang tertulis, mulai dari judul, isi, dan semua teks yang ada dalam artikel. Sedangkan konteks dapat berupa keberadaan artikel tersebut dalam kategori tertentu, atau lainnya.

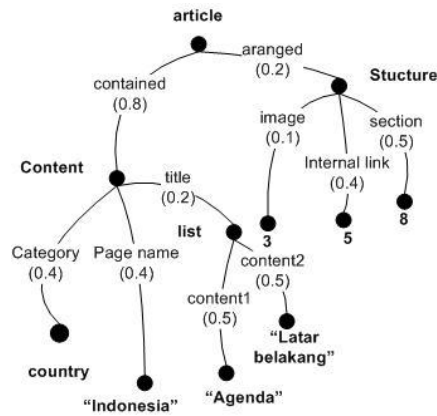
Misalkan seorang pengguna memerlukan informasi mendasar mengenai kluster. Ia memasukkan kata kunci ‘pengenalan’ dan ‘kluster’, dalam konteks Ilmu Komputer, ke mesin pencari semantiknya. Dalam konteks Ilmu Komputer, *term* kluster dapat memiliki dua makna, kluster yang berhubungan dengan Jaringan Komputer atau Analisis Data. Sistem akan menghasilkan dokumen pengenalan kluster berdasarkan jawaban pengguna. Selain itu, mesin pencari tak hanya memberikan dokumen yang mengandung *term* ‘pengenalan’ tapi juga dokumen yang mengandung *term* ‘dasar’ atau ‘tinjauan’.

BAHAN DAN METODE

Metode *Weighted Tree Similarity*

1. Representasi *Tree*

Artikel yang akan dihitung kemiripannya direpresentasikan dalam sebuah *tree* yang memiliki karakteristik node berlabel, cabang berlabel dan cabang berbobot. Gambar 1 menggambarkan representasi *query* pengguna terhadap sebuah artikel Wikipedia. Di dalam representasi artikel tersebut, pengguna mencari sebuah artikel dengan kategori country, nama halaman Indonesia, sub judul agenda dan latar belakang. Artikel memiliki 3 gambar, 5 link dan 8 subbagian. Pada contoh Gambar 1 pengguna lebih menekankan pencari untuk menemukan ketepatan pencarian berdasarkan cabang content (berbobot 0,8) dibandingkan strukturnya (berbobot 0,2). Penentuan tingkat kepentingan cabang ini terdapat di dalam representasi *weighted tree*.



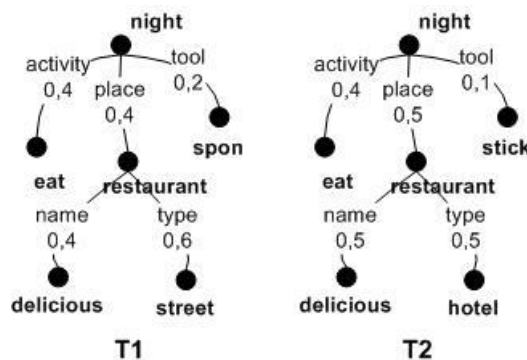
Gambar 1. Contoh Representasi Tree

Representasi tree dalam *weighted tree* mengikuti aturan sebagai berikut (Yang,2005a):

- Nodenya berlabel
- Cabang berlabel
- Cabang berbobot
- Label dan bobot ternormalkan. Label terurut sesuai abjad dari kiri ke kanan. Jumlah bobot dalam cabang setingkat *subtree* yang sama adalah 1.

2. Penghitungan Kemiripan

Algoritma penghitungan kemiripan antara dua *weighted tree* ini terdapat di dalam makalah Bhavsar (2004) dan Yang (2005a). Gambar 2 menunjukkan contoh dua buah *tree* T1 dan T2 yang dihitung kemiripannya.



Gambar 2. Contoh kemiripan penghitungan dasar

Nilai kemiripan tiap pasangan *subtree* berada di antara interval $[0,1]$. Nilai 0 bermakna berbeda sama sekali sedangkan 1 bermakna identik. Kedalaman (*depth*) dan lebar (*breadth*) *tree* tidak dibatasi. Algoritma penghitung kemiripan *tree* secara rekursif menjelajahi tiap pasang *tree* dari atas ke bawah mulai dari kiri ke kanan. Algoritma mulai menghitung kemiripan dari bawah ke atas ketika mencapai *leaf node*. Nilai kemiripan tiap pasang *subtree* di level atas dihitung berdasar kepada kemiripan *subtree* di level bawahnya.

Sewaktu penghitungan, kontribusi bobot cabang juga diperhitungkan. Bobot dirata-rata menggunakan rata-rata aritmatika $(w_i + w'_i)/2$. Nilai rata-rata bobot sebuah cabang dikalikan dengan kemiripan S_i yang diperoleh secara rekursif. Nilai S_i pertama diperoleh berdasar kemiripan *leaf node* dan dapat diatur nilainya menggunakan fungsi $A(S_i)$. Pada awalnya algoritma *weighted tree similarity* hanya memberi nilai 1 bila *leaf node*-nya sama dan 0 bila berbeda (Bhavsar,2004). Perumusan penghitungan kemiripan *tree* ini terdapat di dalam Persamaan berikut:

$$\sum (A(s_i)(w_i+w'_i)^2) \quad (1)$$

dengan $A(s_i)$ adalah nilai kemiripan *leaf node*, w_i dan w'_i adalah bobot pasangan *arc weighted tree*. Penilaian $A(s_i)$ analog dengan logika *AND* sedangkan penilaian bobot pasangan analog dengan logika *OR*.

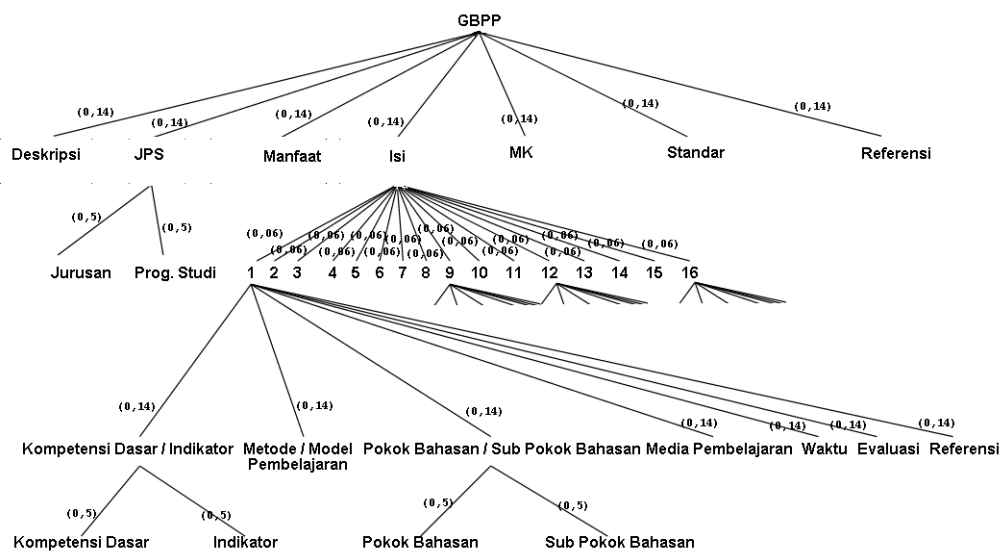
Pada Gambar 2 perilaku algoritma dapat dijelaskan sebagai berikut. Pada awalnya dihitung kemiripan *node* cabang '*activity*' yang diperoleh adalah 1. Kemiripan ini dikalikan rata-rata bobot cabang '*activity*' $(0,4+0,4)/2$ menghasilkan kemiripan cabang. Algoritma kemudian mencari kemiripan *node* cabang berikutnya, '*place*'. Karena *node* ini bukan *leaf* maka algoritma akan turun ke bawah menghitung kemiripan cabang '*name*'. Algoritma kemudian menghitung kemiripan cabang '*type*' dan

diakumulasi dengan kemiripan cabang 'name'. Akumulasi ini merupakan nilai kemiripan subtree 'restaurant'. Algoritma bergerak ke cabang tool dan akumulasi kemiripan dengan cabang lain yang setingkat menghasilkan kemiripan total.

HASIL DAN PEMBAHASAN

Pohon berbobot untuk pencocokan dokumen dibuat dengan cabang tanpa label, hanya ada nilai bobot pada cabang. Besarnya nilai bobot dibagi sama rata berdasarkan jumlah cabang, dengan nilai antara mulai dari 0 sampai dengan 1.

Contoh kata-kata kunci yang digunakan untuk pencocokan dokumen GBPP, diantaranya: GBPP/SILABUS, Jurusan, Program Studi, MataKuliah/SKS/Kopel, Deskripsi Singkat, Standar Kompetensi, Manfaat MataKuliah, Kompetensi Dasar, Indikator, Pokok Bahasan, Sub Pokok Bahasan, Metode/Model, Media Pembelajaran, Waktu, Referensi. Pada pohon berbobot GBPP terdapat tujuh cabang utama dengan bobot masing-masing cabang adalah 0,14 dan memiliki kedalaman 4 tingkat. Pohon berbobot untuk pencocokan dokumen GBPP dijelaskan pada gambar 3.



Gambar 3. Pohon GBPP



KESIMPULAN DAN SARAN

Telah dibuat pohon berbobot sebagai pembanding untuk pencocokan dokumen perkuliahan yang nantinya akan diterapkan menggunakan algoritma *weighted tree similarity*. Dokumen perkuliahan yang dimaksud adalah GBPP, SAP dan kontrak perkuliahan.

Pohon pembanding dibuat dengan cabang tanpa label, hanya ada nilai bobot pada cabang. Besarnya nilai bobot pada tiap cabang dibagi sama rata berdasarkan jumlah cabang, nilainya antara 0 sampai dengan 1.

DAFTAR PUSTAKA

- Bhavsar VC, Boley H, Yang L. 2004. A weighted tree similarity algorithm for multiagent system in e-business environments. *Computational Intelligence* 20(4): 584-602.
- Borang Monev Proses Perkuliahan Bermutu, SOPMO2. 2013. Pusat Penjamin Mutu Universitas Lampung, Bandar Lampung.
- Oxford. 1994. *Pocket Oxford Dictionary*. Oxford University Press. Oxford.
- Prosedur Mutu Proses Perencanaan Perkuliahan, PM-9, 2014. Jurusan Ilmu Komputer FMIPA Universitas Lampung, Bandar Lampung.
- Rahutomo F. 2008. Penerapan Algoritma Weighted Tree Similarity Untuk Pencarian Semantik. *JUTI*. 7(1): 39-46.
- Wikipedia. 2015. Semantics. <http://en.wikipedia.org/wiki/Semantics>.
- Yang L, Sarker BK, Bhavsar VC, Boley H. 2005. A weighted tree simplicity algorithm for similarity matching of partial product descriptions. Proceeding of ISCA 14th International Conference on Intelligent and Adaptive Systems and Software Engineering. Toronto Canada. pp. 55-60